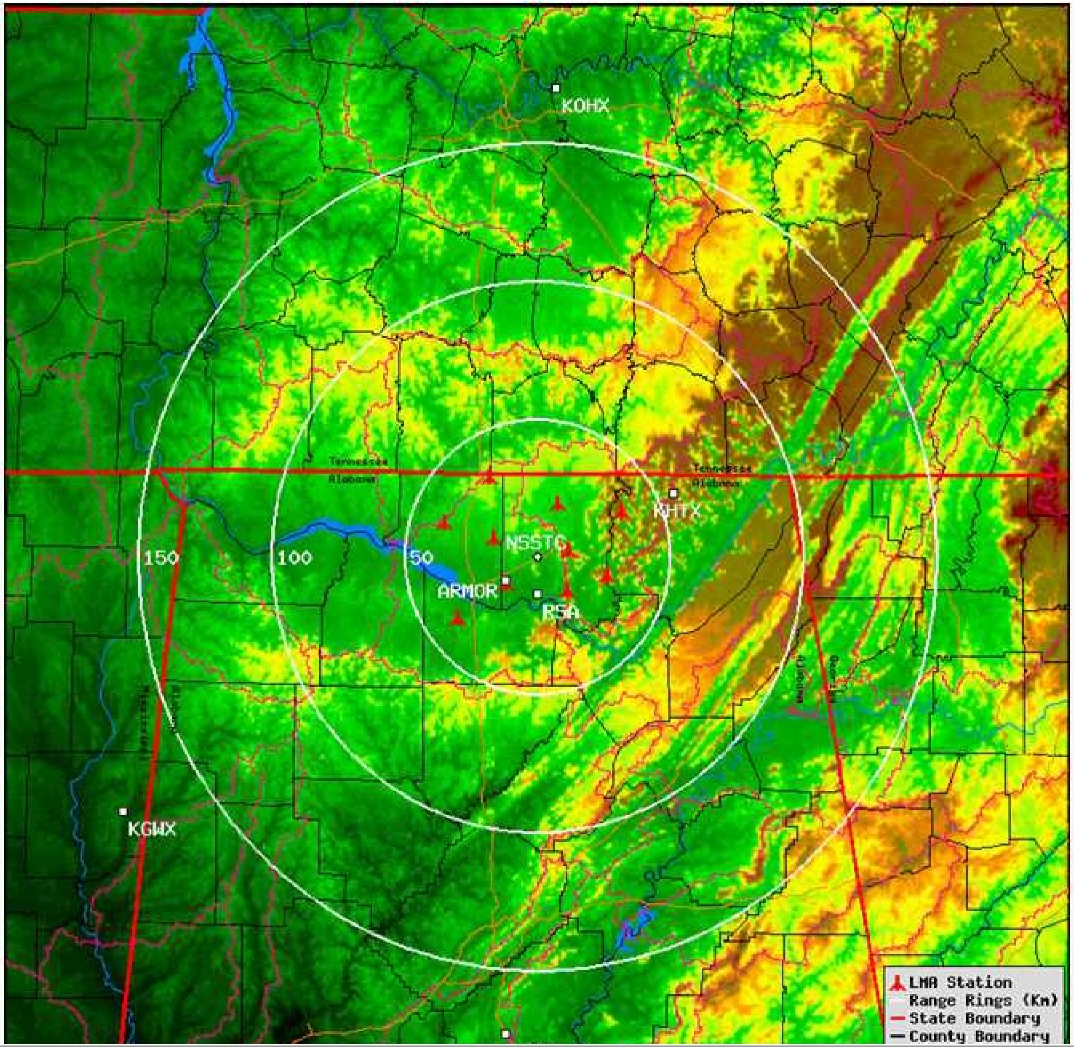
A Flash Clustering Algorithm for North Alabama Lightning Mapping Array Data

E. W. McCaul, Jr., USRA Huntsville
J. E. Bailey, Raytheon ITSS
J. Hall, SAIC
S. Goodman, R. Blakeslee, NASA MSFC
D. E. Buechler, UAH

Background

- LMA systems use time of arrival of VHF radiation from LTG channel segments to infer time, location of sources
- LMAs are prone to noise of several types
- Accurate flash clustering and noise removal are crucial for meaningful use of data
- Source counts depend strongly on r, while flash counts do not

North Alabama LMA With Digital Elevation Map



LMA flash algorithm: methodology

- Use standard time, space clustering criteria of 0.3 sec, 1 km + r-squared location error to cluster sources into flashes
- Impose no artificial upper limits on duration of flashes
- Identify and remove non-meteorological events so that storms can be studied

LMA flash algorithm: comments

- We find "singleton" flashes comprise a significant proportion of our data (often more than 10%)
- Some of the singletons are noise, but many appear to be meteorological
- We find a few long-duration events (5-10+ s) associated with complex "chain-reaction" flashes; these may contain multiple CG points

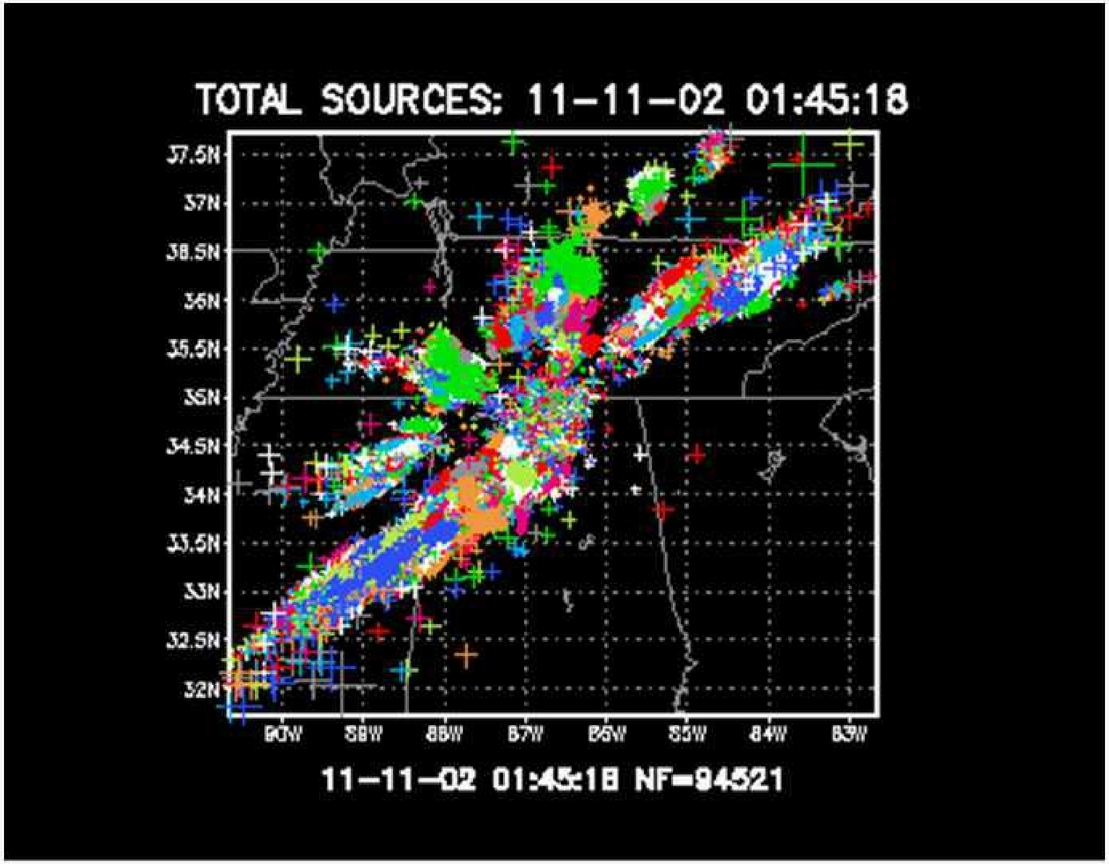
Types of singletons

- 1. Meteorological events (mostly in anvils)
- 2. Network clutter at close range
- 3. Rarely, fixed site corona(?) during stormy weather
- 4. Interference from distant TV ch. 5 signals
- 5. Aircraft tracks in ice/mixed clouds
- 6. Misplaced or mistimed sources (COMMON and hard to remedy!)

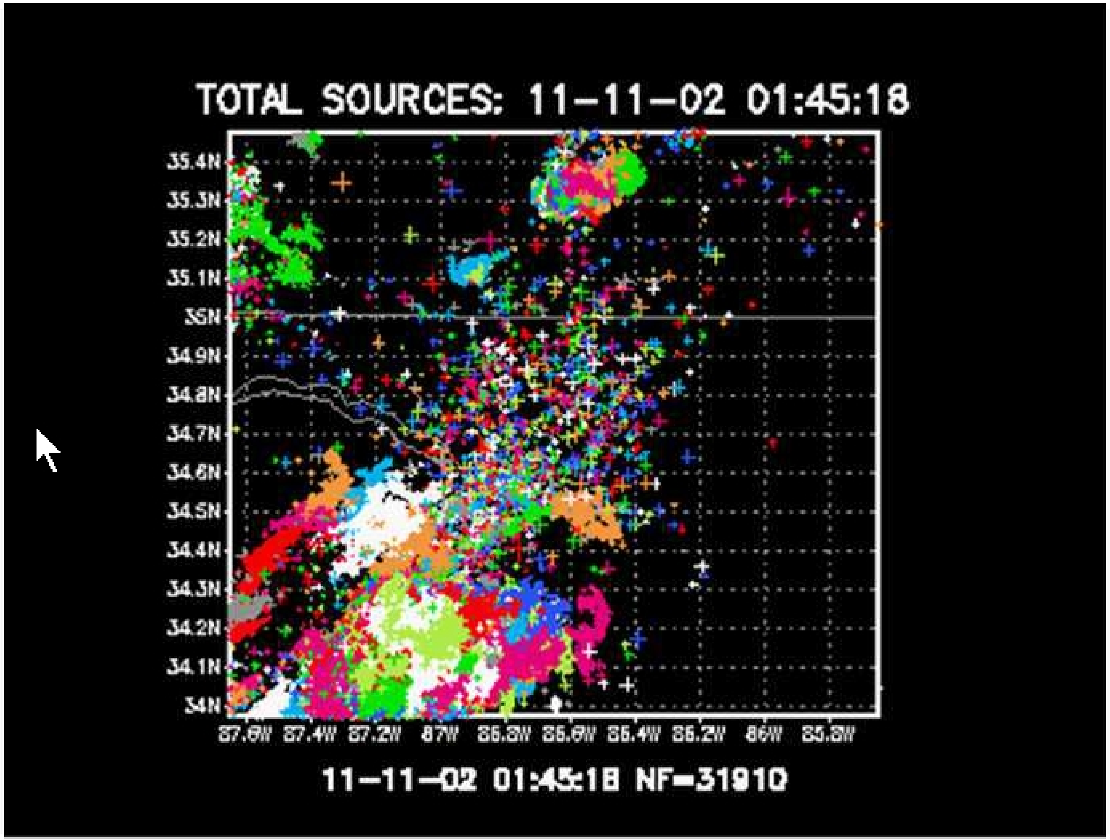
LMA flash algorithm: methodology

- 1. Generate preliminary flash lists based on the space and time cluster criteria
- 2. Identify, remove aircraft tracks
- 3. Identify, remove network clutter
- 4. Identify, remove fixed-site corona
- 5. TV ch. 5 interference is sparse; ignore
- 6. Remainder of events are weather

Consider a sample from 11 Nov 2002, 0145 UTC



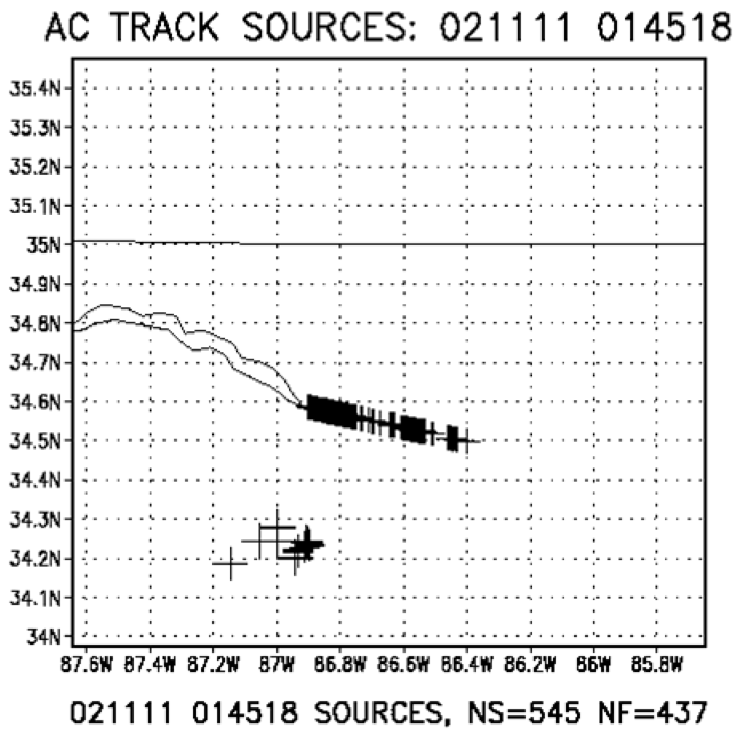
Now let's zoom in on the LMA local area:



Identification of AC tracks

- Exploit slow movement of AC relative to real lightning to study AC tracks
- Treat AC tracks as "slowly propagating LTG events"
- Search all flashes with n<6 events, check for propagation vel < 350 m/s, except use 3000 m/s if dt < 0.1 s; these velocities are at least 3-4 orders of magnitude slower than real LTG
- this identifies most AC track sources, but a few sources in active storms may also be falsely selected

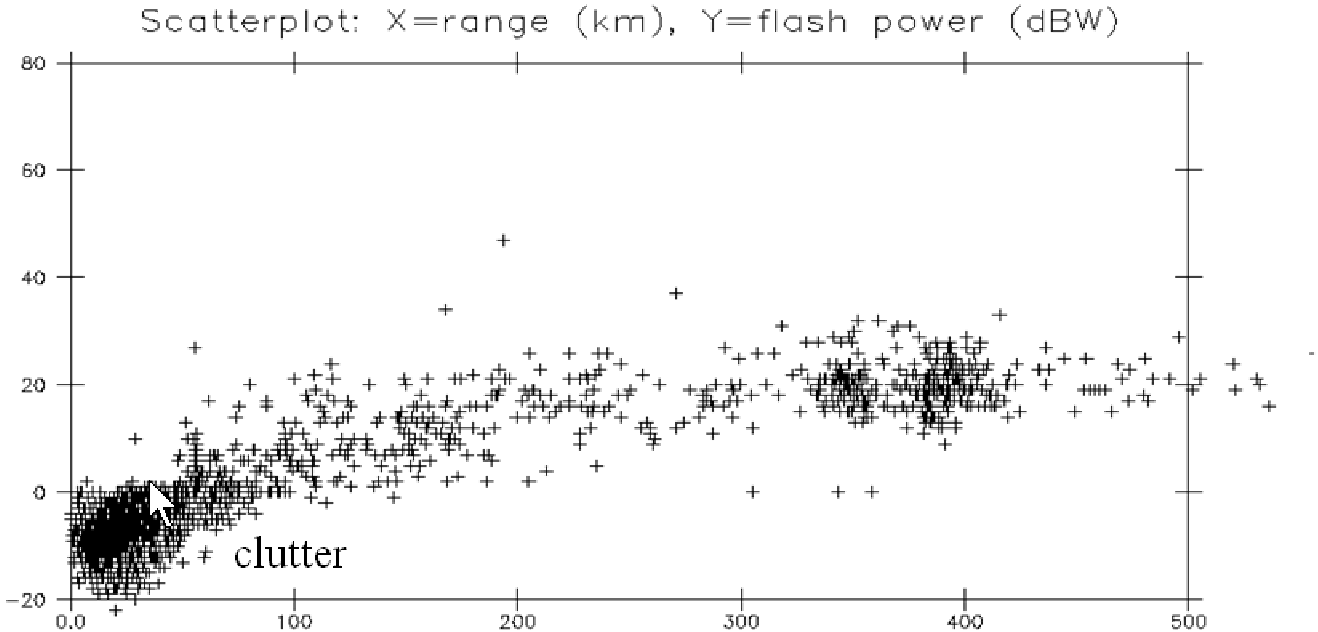
We find these AC track sources:



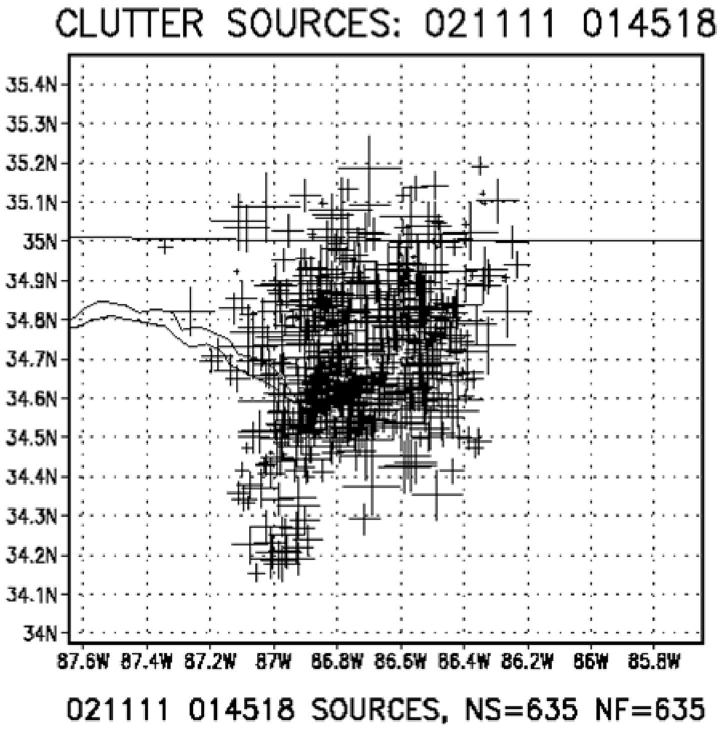
Identification of network clutter

- Look for other singletons having:
- Power < 2 dBW,
- Range < 90 km from LMA centroid
- These are considered to be clutter noise

Scatterplot of power vs range shows clutter nicely:



We find this "clutter" pattern:



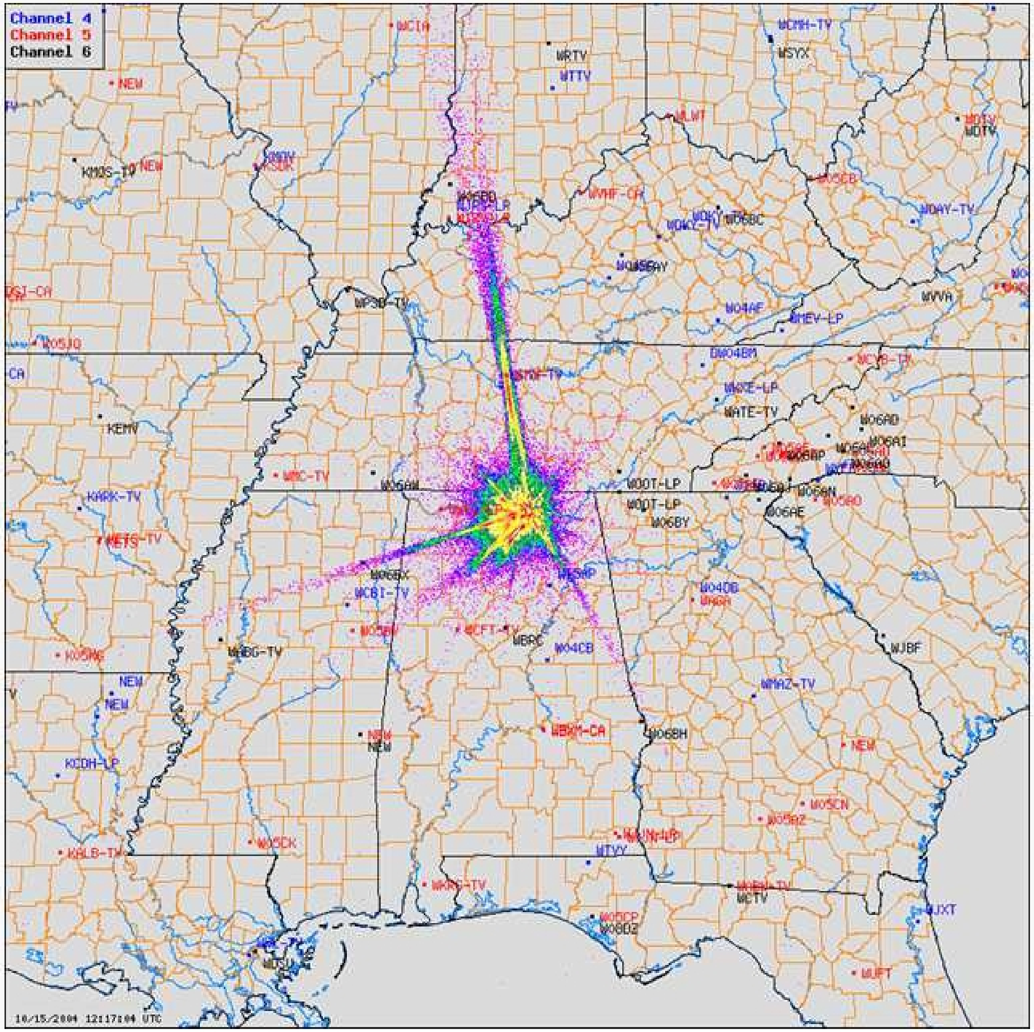
Identification of fixed-site corona

- We have found three sites that produce sporadic bursts of sources during storms; at least one has been found to be a home radio antenna
- Check data for any sources that:
- Are very close to known source x,y
- Are identified at altitudes < 3 km
- These are fixed-site corona events

Identification of TV ch. 5 noise

- Distant TV ch. 5 signals produce sources all along four known radials subtending the TV transmitters
- These signals occur at < 1 per minute, even less often during stormy weather
- Search for them is optional, but is usually safely neglected for our ~5 min analyses of storms

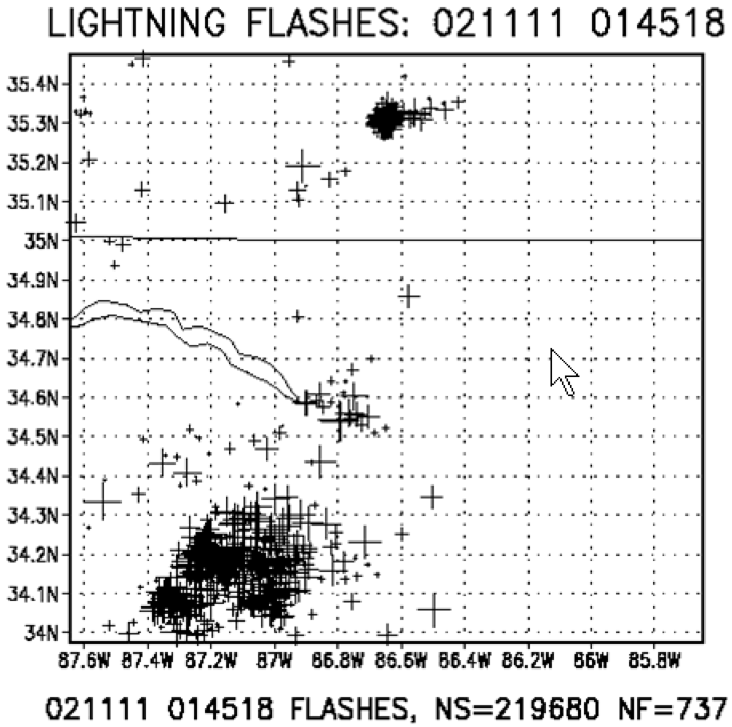
Long integration shows TV Ch. 5 noise:



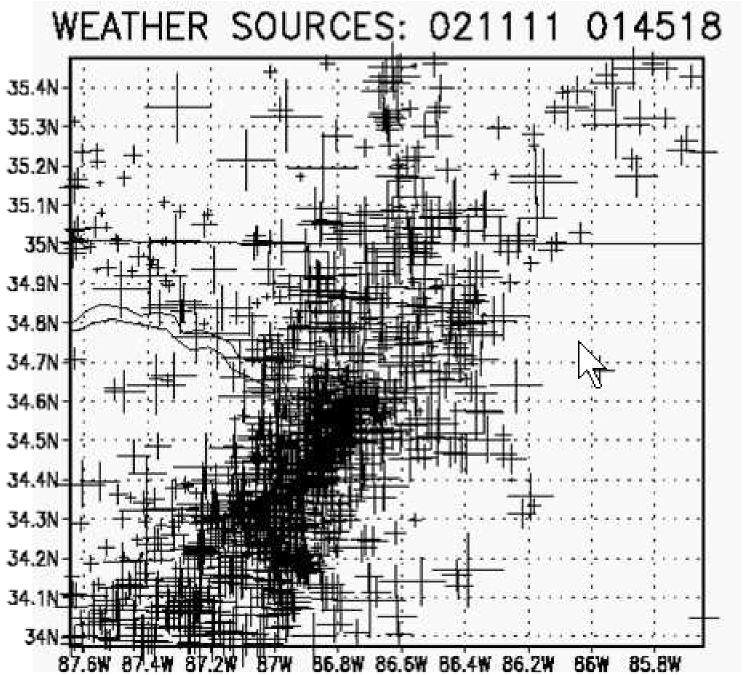
LMA flash algorithm: methodology

- If a singleton is not AC track, not network clutter, not corona, and not TV noise, then we assume it is meteorological
- If a multiple-source event is not an AC track, then we assume it is meteorological

Starting points of multiples:



…and meteorological singletons
(mostly in storm anvils):



021111 014518 SOURCES, NS=1146 NF=1146

Conclusions 1

- Met. multiple-source flashes show large variations in x,y,z,t structure
- Met. singletons correlate with met. multiples
- Clutter patterns vary with time, but show little spatial correlation with weather
- AC tracks obviously vary with amounts of flight activity into ice/mixed cloud
- AC tracks, clutter can be identified, removed with reasonable accuracy; some false positives occur, which may be removed using source density

Conclusions 2

- Biggest contributor to noise is mistimed or mislocated sources; this leads to large spread along range of distant storm source clusters
- Our radial range location error accounts for this; however, beyond 160 km, this uncertainty rivals typical cell spacing
- We avoid doing spatial analyses of cells for ranges beyond ~150 km